

A Heuristic Method for Complement Common Visual Localization

충남대학교 메카트로닉스공학과 영상시스템연구실 김태원

2020. 8. 3.

1. 기술의 특징

제안된 방법은 Image retrieval과 keypoint matching을 사용한 common visual localization method와 Stereo Direct Sparse Odometry(Stereo DSO)를 통해 복원한 test set의 3d point cloud의 위치를 train set의 전체 lidar map에서 찾는 heuristic method로 구성된다. 따라서 train set에 일치하는 사진이 없거나 두 사진에서 keypoint matching이 성공하지 않는 경우에도 위치를 찾을 수 있다.

2. 구체적인 구현 방법

A. Visual localization method

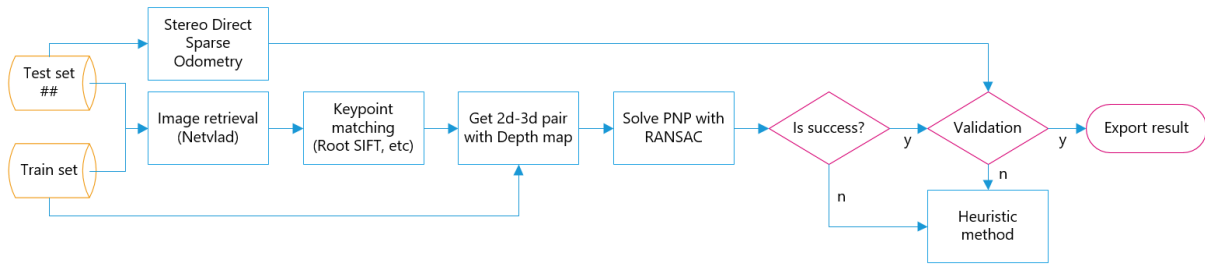


그림 1. Flow chart of visual localization method

Visual localization method는 그림1과 같이 Netvlad를 통해 획득한 train set과 test set의 image pair(I^{test}, I^{train})에서 Root-SIFT, VGG, DAISY, ORB 등 여러 keypoint matching 알고리즘을 혼합하여 2d-2d point pair(p_i^{test}, p_i^{train})를 구한다[1][2][3]. 그 다음 I^{train} 이 촬영 된 시간에서 ± 30 초 이내의 lidar data들을 I^{train} 에 투영하여 depth map D 를 만든다. D 를 이용하면 2d-3d point pair(p_i^{test}, p_i^{train})를 구할 수 있는데 D 는 sparse 하므로 일부 p_i^{train} 에 대한 p_i^{train} 은 바로 구할 수 없다. 이 경우, D 에서 p_i^{train} 을 기준으로 16pixel 거리 내의 3d point들을 모아 평면을 구성하고 I^{train} 의 카메라 원점과 image plane의 p_i^{train} 을 지나는 3차원 viewing ray와의 교점을 구하여 사용하였다.

2d-3d point pair를 구했기 때문에 PNP 문제를 풀어서 I^{test} 와 I^{train} 사이의 pose를 구할 수 있다. 만약 solve PNP를 위한 충분한 수의 2d-3d point pair가 구해지지 않았거나 solve PNP가 실패하면 이후에 기술 할 heuristic method로 넘어간다. 만약 solve PNP 과정이 성공하여 pose가 구해지면 이를 한번 더 검증한다. I^{train} 의 pose와 I^{test} 의 pose는 비슷한 방향을 바라봐야 하며, 특히 차량에 설치 된 센서이기 때문에 global coordinates에서 z축 높이가 비슷해야 한다. 추가적으로 Stereo

DSO를 통해 획득한 test set의 point cloud의 각 포인트와 I^{train} 근처의 lidar cloud 사이의 nearest Euclidean distance의 평균을 판단의 근거로 사용하였다.

그림 2는 제안한 visual localization method의 결과를 나타낸다. 상단의 두 이미지는 I^{test} 와 I^{train} 을 나타내며 keypoint matching을 통해 구한 2d-2d point pair를 표시 한 것이다. 하단의 point cloud map은 I^{train} 근처의 lidar data 들을 나타내며 주변보다 밝게 표시된 점들은 Stereo DSO를 통해 획득한 point cloud를 solve PNP를 통해 구한 pose로 transform하여 표시한 것이다. 중단의 두 이미지는 임의의 좌표축을 양쪽 이미지에 re-projection 하여 결과를 확인한 것이다.

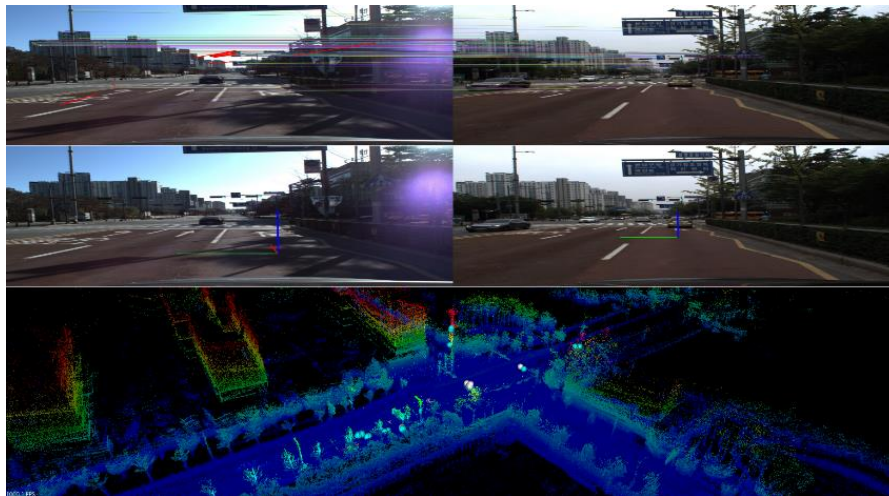


그림 2. Result of visual localization method

B. Heuristic method

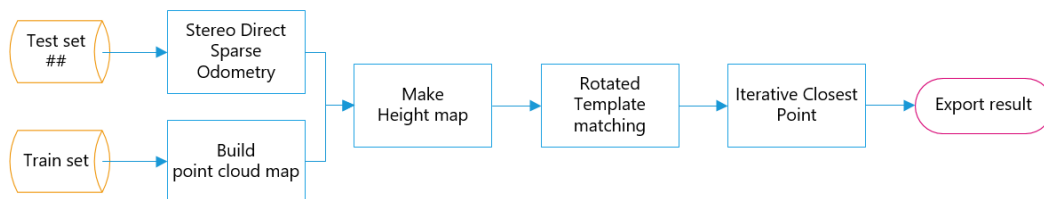


그림 3. Flow chart of heuristic method

Stereo DSO는 Direct Sparse Odometry(DSO)를 stereo vision에서 사용 할 수 있게 확장한 것이다 [4]. 따라서 test set으로 주어진 50장의 stereo image sequence에서 point cloud를 획득 할 수 있다. 이때, 각 test set의 마지막 image의 local axis를 기준으로 point cloud를 정리한다. 그림 4는 Stereo DSO로 구한 point cloud 들을 나타낸다. 본 연구에서는 Stereo DSO의 parameter를 조정하여 속도를 희생하는 대신 accuracy와 point cloud의 density를 높이도록 하였다.

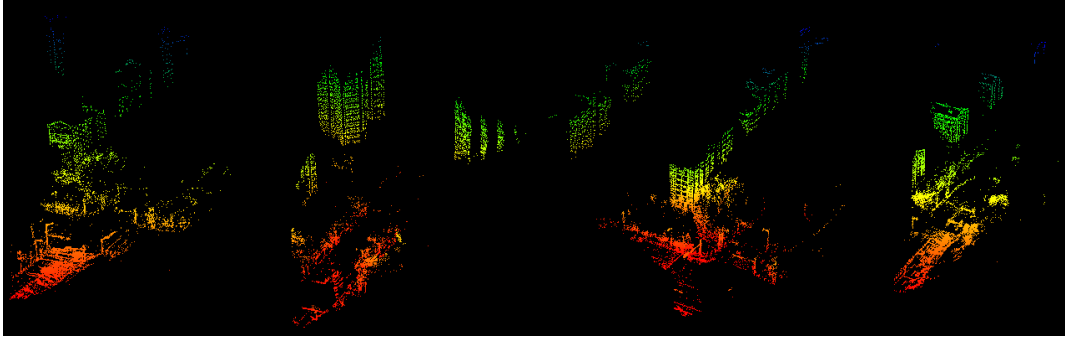


그림 4. Point cloud of test sets using Stereo DSO

Global point cloud map은 train set의 lidar data와 ground truth pose들을 이용하여 만든다. 이후 point cloud는 height map의 형태로 2d grid에 놓여지며 x, y축 sampling은 1m마다 수행하였다. 같은 grid에 여러 point가 중복되는 경우 z값이 큰 것을 우선하였다. Test set의 point cloud도 같은 방법을 사용하여 height map을 만든다.



그림 5. Result of build height map, left: pangyo train set(crop and resized), right: test set pangyo05

그림 5와 같이 두 height map이 준비 되었다면 일반적인 template matching과 같은 방법으로 test set의 height map을 train set의 height map에 순회 시키며 SAD 값을 구한다. 이때 회전 각을 구하기 위해 test set의 height map을 10도 간격으로 회전시켜 SAD 값이 최소가 되는 3DOF pose(x, y, yaw)를 구한다. 그 후, Global point cloud map과, extrinsic parameter를 곱해 vehicle 좌표축으로 변환한 test set의 point cloud를 입력으로 ICP method를 사용하여 6DOF pose를 구한다.

ICP의 initial guess는 앞에서 구한 3DOF pose와 train set의 pose들 중 3DOF pose와 가장 가까운 pose에서 z값을 추출하여 넣어준다. 그림 6은 visual localization이 실패한 pangyo05와 pangyo17에서 제안한 heuristic method를 사용하여 위치를 찾은 결과이다.

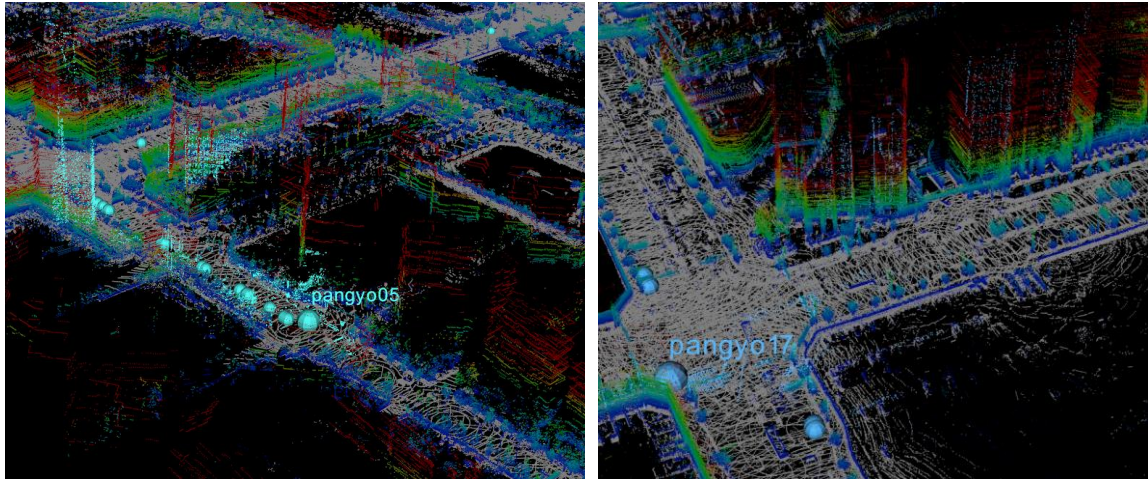


그림 6. Result of heuristic method for pangyo05, pangyo17

Reference

- [1] Arandjelovic, Relja, et al. "NetVLAD: CNN architecture for weakly supervised place recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [2] Arandjelović, Relja, and Andrew Zisserman. "Three things everyone should know to improve object retrieval." *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012.
- [3] Simonyan, Karen, Andrea Vedaldi, and Andrew Zisserman. "Learning local feature descriptors using convex optimisation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.8 (2014): 1573-1585.
- [4] Wang, Rui, Martin Schworer, and Daniel Cremers. "Stereo DSO: Large-scale direct sparse visual odometry with stereo cameras." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.